Spatial Variability and Robust Interpolation of Seafloor Sediment Properties Using the SEABED Databases

John A. Goff
Institute for Geophysics, Jackson School of Geosciences, University of Texas
JJ Pickle Research Campus
10100 Burnet Rd. (R2200), Bldg. 196, Austin, TX 78758
phone: 512-471-0476 fax: 512-471-0999 e-mail: goff@ig.utexas.edu

Award Number: N00014-05-1-0079

LONG-TERM GOALS

This project is a collaborative effort with C. Jenkins at the Univ. Colorado. The long-term goals are to: (1) advance the understanding of the spatial variability of seabed properties as a function of geologic environment; (2) develop robust means of interpolation in the presence of uncertain data, (3) provide for the estimation of uncertainty in the interpolation at unsampled locations, and enable investigation of optimal survey design to minimize uncertainties; and (4) publish a computational/database structure capable of producing seafloor maps of wide geographic extent, for multidisciplinary use - in global change issues, defense, engineering, and ecology.

OBJECTIVES

The US Geological Survey, in collaboration with Chris Jenkins of INSTAAR/Univ. Colorado, has recently released a large data base of seabed sedimentary properties in US coastal and shelfal waters. Dubbed "usSEABED" (Figure 1; Williams, et al., 2003; Reid et al., 2005, 2006; Buczkowski et al., 2006; http://walrus.wr.usgs.gov/usseabed; USGS publications DS 118 (Atlantic), 146 (Gulf of Mexico) and 182 (Pacific)), the data base is a compilation of available records of sedimentary data from seafloor samples, cores, and visual observations. Over 120,000 independent seafloor mean grain size measurements are included, along with many other data types. This work is part of a larger effort to develop a world-wide data base (collectively, the "dbSEABED" data bases), which continues to enlarge every year with additional records (http://instaar.colorado.edu/~jenkinsc/dbseabed).

The comprehensive dbSEABED data bases provide a new and unprecedented opportunity for advancing Navy interests related to the acoustic response of the seafloor: (1) they will enable the creation of maps of seafloor sedimentary properties over areas and at a level of detail previously unobtainable with single campaign efforts; and (2) they will enable investigation of the variability of sedimentary properties, which is a critical factor in assessing uncertainty in acoustic detection, over a wide range of environmental conditions. However, the highly heterogeneous nature of these records present important challenges in data handling. The most significant issue involves the measurements of mean grain size either by "extracted" or "parsed" methods. Extracted measurements are derived from analytic methods of computing the grain size histogram, such as sieves, settling tubes or diffractometry. Such measurements tend to be very precise. Parsed measurements of mean grain size are derived from a calibrated conversion of a word-based description of sediments (e.g., sand, fine sand, mud, muddy sand, gravelly sand, etc.). Although less precise, such measurements provide the only comprehensive coverage in many regions. The extent to which these two forms of measuring

Report Documentation Page					*Form Approved* *OMB No. 0704-0188*
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE **2007**		2. REPORT TYPE			3. DATES COVERED **00-00-2007 to 00-00-2007**
4. TITLE AND SUBTITLE **Spatial Variability and Robust Interpolation of Seafloor Sediment Properties Using the SEABED Databases**					5a. CONTRACT NUMBER
					5b. GRANT NUMBER
					5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)					5d. PROJECT NUMBER
					5e. TASK NUMBER
					5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of Texas,Institute for Geophysics, Jackson School of Geosciences,10100 Burnet Rd. (R2200), Bldg. 196,Austin,TX,78758**					8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)					10. SPONSOR/MONITOR'S ACRONYM(S)
					11. SPONSOR/MONITOR'S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**11**	

mean grain size are compatible is a critical outstanding and a consequent objective of our work. Observational bias represents a second significant issue. Bias is most prominently manifest in the inclusion or exclusion of coarse fraction (e.g., shell hash), depending on the needs of the observer. Extracted measurements appear more likely to exclude shells, which are important to acoustic considerations.

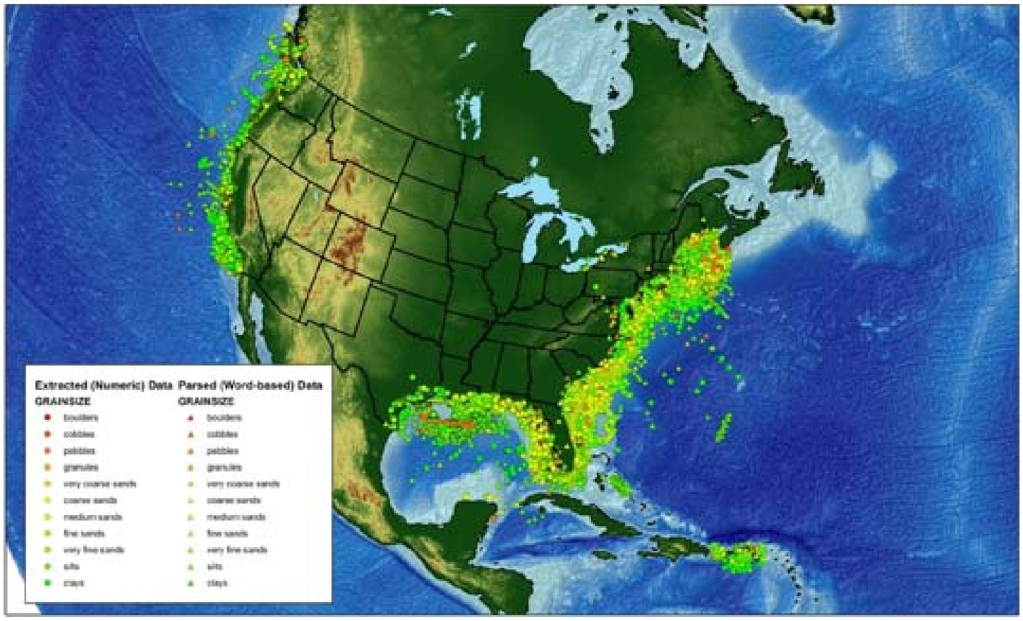

Figure 1. Location map of current (2007) usSEABED data coverage (250,000 records), color coded by mean grain size (Williams, et al., 2006).

APPROACH

We formulated a methodology for statistical analysis of randomly-located marine sediment point data, and applied it to the U.S. continental shelf portions of usSEABED mean grain size records. The usSEABED database, like many modern, large environmental datasets, is heterogeneous and interdisciplinary. We statistically test the database as a source of mean grain size data, and from it provide a first examination of regional seafloor sediment variability across the entire US continental shelf. Data derived from laboratory analyses (“extracted”) and from word-based descriptions (“parsed”) are treated separately, and they are compared statistically and deterministically. Data records are selected for spatial analysis by their location within sample regions: polygonal areas defined in ArcGIS chosen by geography, water depth, and data sufficiency (Figure 2). We derive isotropic, binned semivariograms from the data, and invert these for estimates of noise variance, field variance, and decorrelation distance (Figure 3). The highly erratic nature of the semivariograms is a result both of the random locations of the data and of the high level of data uncertainty (noise). This decorrelates the data covariance matrix for the inversion, and largely prevents robust estimation of the fractal dimension.

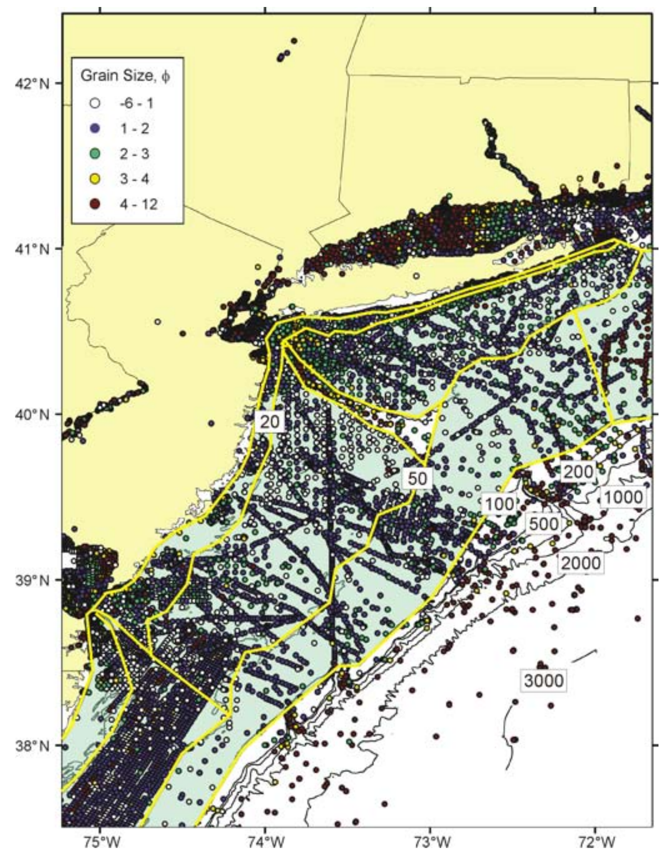

Figure 2. Location of usSEABED records within the mid-Atlantic Bight, color coded by mean grain size, and overlain on bathymetric contours (meters). Sample areas defined for this region are indicated by green polygons with yellow boarders.

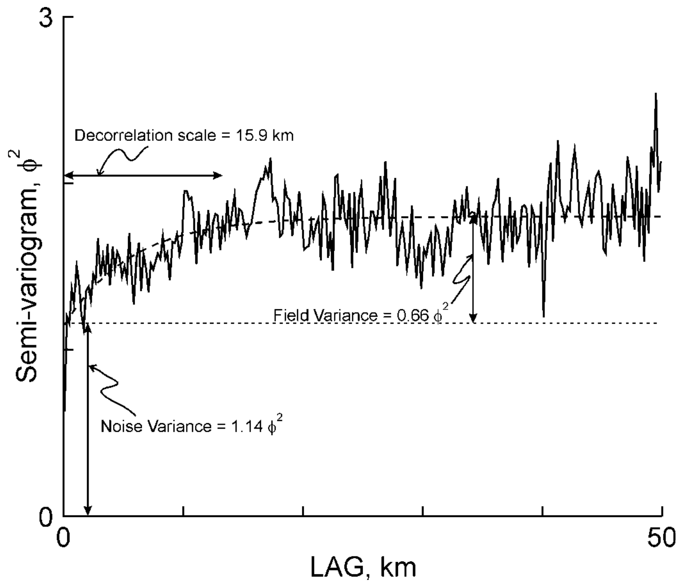

Figure 3. A binned semivariogram (solid) derived from parsed mean grain size measurements in the 0-20 m depth range of the New York Bight (Figure 2). The best-fit von Kármán model with noise spike is overlain (dashed), with parameter values as indicated. The fractal dimension of the model is 1.5, which corresponds to an exponential curve.

WORK COMPLETED

Project funding is now complete. Goff's primary tasks for this project were (1) developing a tool for correcting noisy data through resampling and (2) utilizing the usSEABED database to explore seabed variability in US shelfal waters. The first task was completed and published in Goff et al. (2006). The second task was completed last year and is the topic of an article under review (Goff et al., in review). This progress report will focus on the latter task, for which a preliminary account was given in last year's annual progress report.

RESULTS

As primary component of the study we examined the suitability of the aggregated usSEABED data collection for mapping and variability analysis. Our quantitative comparison (Figure 4) between the parsed and extracted forms of mean grain size data reveal some differences. As expected, the noise variance tends to be larger for the parsed records (by ~0.2-1.0 ϕ^2), which reflects a higher level of uncertainty in the measurements. Greater temporal variability (i.e., timing of sample collection) may also be important. At present, temporal information cannot be extracted from the usSEABED database, but it is likely that the word-based data records span a much greater range in dates. Any temporal effects on grain size measurements (e.g., changes in sedimentary conditions, changes in navigational resolution) will presumably factor into the data uncertainty. Higher levels of uncertainty in the parsed measurements might also be related to the likelihood that they are more likely to incorporate a wider set of materials, such as shells.

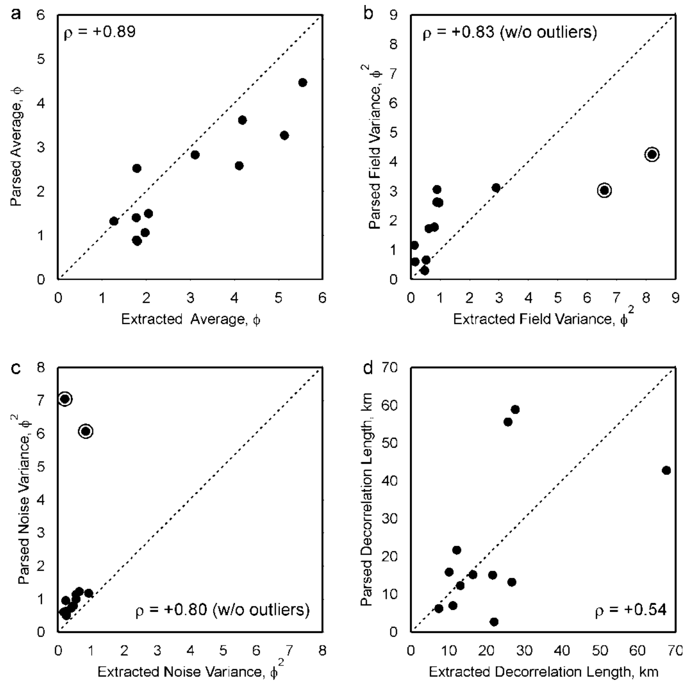

Figure 4. Plots of extracted versus parsed statistical parameters from sample areas with adequate coverage of both types of data records. Dashed line indicates 1:1 correspondence. Circled symbols in (b) and (c) are from Gulf of Maine sample regions, and are discussed in the text. Correlation coefficients (ρ) are given for each plot, neglecting the outliers.

In general, the extracted mean grain sizes tend to exhibit higher ϕ values (finer grain sizes), ~0.5 ϕ on average, and lower field variance relative to the parsed mean grain sizes. Both observations might be explained by a tendency for grain size analysts to discard the very coarsest fraction of a sediment, particularly if it contains shell material. These differences between parsed and extracted measurements are, however, somewhat regionally dependent, and it is not possible to formulate a precise universal conversion factor between the two. Nevertheless, if sufficient numbers of each type of data exist within a particular sample region, it should be possible to empirically define a local conversion so that the two types of data can be used together, along with their respective uncertainties, for quantitative applications such as mapping.

Our analysis of sample regions for the usSEABED records of mean grain size on the continental shelf reveal considerable geographic variability in the estimated parameters of average (Figure 5) field variance (Figure 6) and decorrelation distance (Figure 7). High field variances and short decorrelation lengths on the Florida shelves may indicate a high level of patchiness due to shelly material. Very high variances in the Gulf of Maine may be a result of residual glacio-fluvial gravel patches interspersed with fine-grained sediments. Elsewhere, we observe a fairly strong inverse relationship between the average mean grain size and the field variance (expressed as a positive correlation in ϕ units). We are uncertain as to the cause of this correlation.

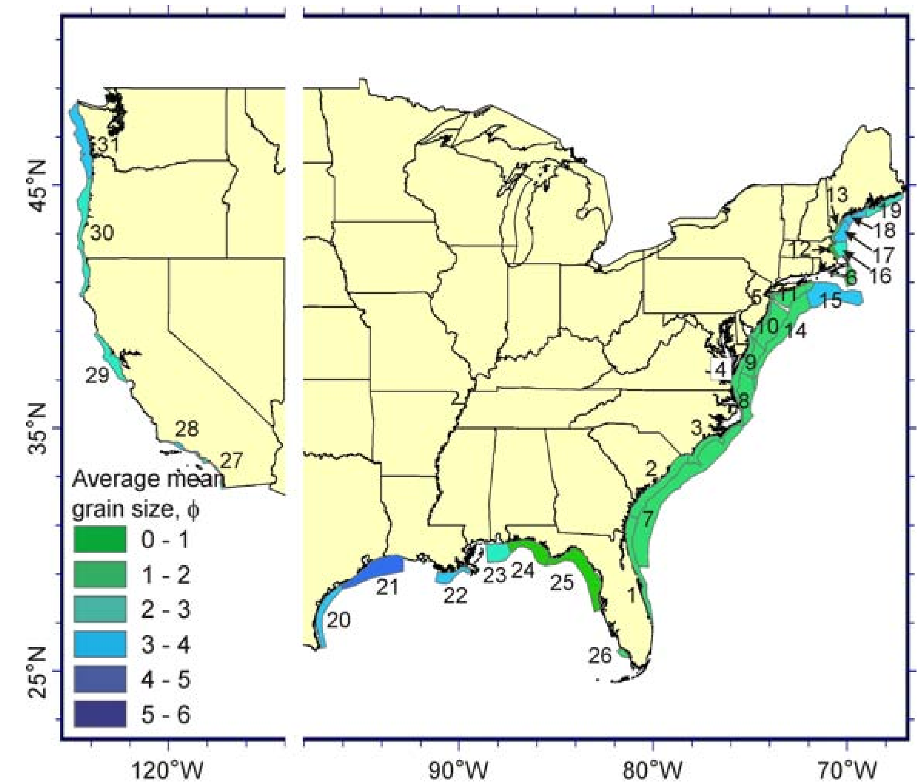

Figure 5. Sample areas defined over the entire usSEABED database, color coded by the average of the parsed mean grain size measurements within the sample area.

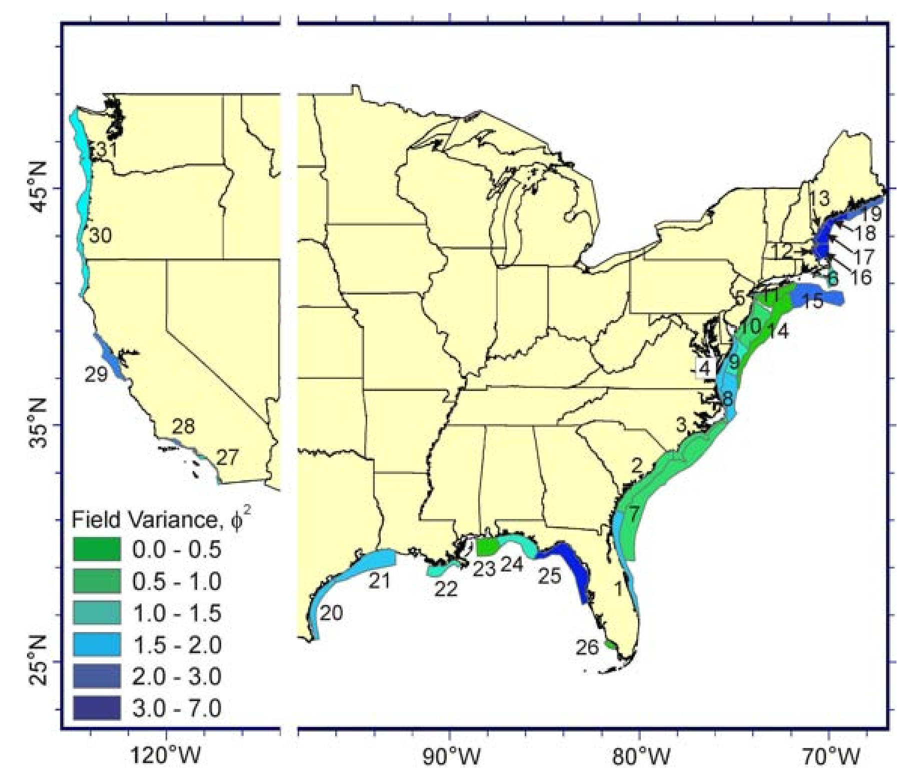

Figure 6. Sample areas color coded by estimated field variance of mean grain size measurements.

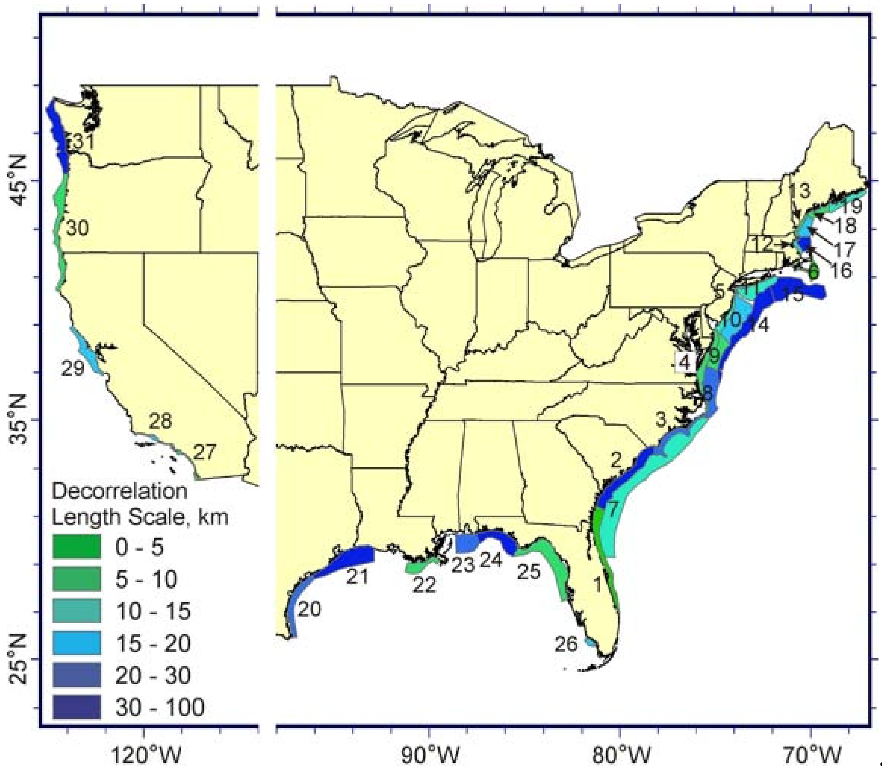

Figure 7. Sample areas color coded by estimated decorrelation distance of mean grain size measurements.

Other than the small values on the Florida shelf, the estimated decorrelation length scales do not present coherent geographical relationships. Unlike the results of Jenkins and Goff (in review) for the analysis of mean grain size measurements in the Adriatic Sea, we do not find evidence on the U.S. Atlantic margin for any consistent depth relationship for this parameter (other regions were insufficiently sampled to discriminate sample regions based on water depth). We believe that analysis of more sample areas from a greater variety of settings will be needed to decipher the primary influences on decorrelation length scale. We suggest here that it may be controlled by competing relationships of geologic inputs (e.g., sediment facies), which probably tend toward larger decorrelation length scales, and oceanographic reworking, which probably tends toward shorter length scales (e.g., bedforms).

Comparison of extracted versus parsed statistical parameters provides evidence that the noise variance estimated from parsed and extracted mean grain size measurements are correlated. Assuming the noise variance is related only to the data uncertainty, there is no reason to expect such a correlation, suggesting that noise is somehow influenced by the properties of the field. However, no such evidence could be found in our interparameter comparisons. To explain these observations, we hypothesize that a very short-scale of field variability exists that is superimposed on the larger scale of variability that we discern through estimate of the decorrelation length of the semivariogram, and that the decorrelation length of this shorter scale variability is shorter than the resolution scale of the sample data. In other words, the portion of data variability that we identify as “noise” includes both a real field component and a data uncertainty component. If true, then we cannot directly distinguish between the two, although we may be able to infer the field component if we are able to postulate globally constant values of uncertainty for parsed and extracted measurements. More data analysis will be required to determine if that is the case.

Our example using the Long Island shelf data (Figure 8) shows that usSEABED can reliably be utilized for creating maps of seafloor mean grain size and possibly other sediment characteristics. Due to the noisy character of the data, some sort of filtering or other noise reduction algorithm (e.g., Goff et al., 2006) is recommended prior to map generation. To combine the parsed and extracted measurements, a bias correction proxy must be applied, and such a correction should be evaluated individually for each region of interest. For the Long Island shelf data, a simple mean correction of 0.5 ϕ was found to be suitable because the semivariogram statistics were otherwise found to be very similar between the two types of data. Recognizing that coarse content is excluded from many analytic results, we applied the correction by subtracting it from the extracted data. Other regions, however, exhibit significant difference in both the mean and variance of parsed versus extracted mean grain size values, and in those cases a more complex bias correction proxy must be devised.

IMPACT/APPLICATIONS

This project could provide a major advance in marine science, a set of reliable methods which transform point-site seabed data into griddings that will be useful across oceanographic disciplines, sediment transport, acoustics, habitat, wave-energy generation. Our work will result in a set of software tools that will be open source, and available for inclusion any existing software packages. These tools could be of importance to the Navy, particularly in dealing with areas with sparse data, such as “denied” areas. In particular, an understanding of the relationship between environmental parameters, geologic setting and spatial variability could provide an ability to predict the amount and spatial scales of seabed variability using a parameterized semi-variogram model. This functionality provides a basis upon which to predict seabed parameters at unsampled locations, and to assess the uncertainty in that prediction. Such an understanding will have important implications for assessing acoustic prediction uncertainty. Furthermore, the semi-variogram model can be used to investigate optimal survey design, should it be possible to conduct limited sampling in denied areas via covert means (e.g., AUV’s).

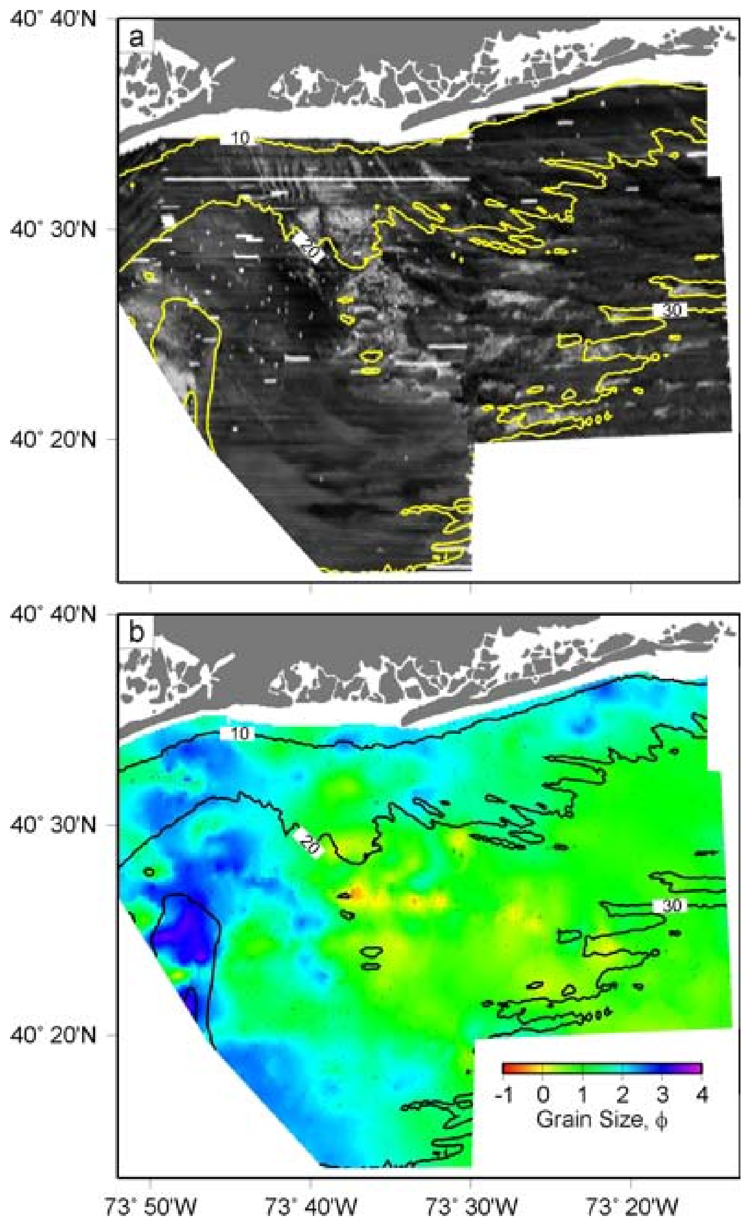

Figure 8. Comparison of overlapping portions of USGS acoustic backscatter data (a; from Schwab et al., 2000) and the interpolated, resampled mean grain size off the western Long Island shelf (b). Lighter shades indicate in (a) higher backscatter and coarser sediment, darker shades indicate lower backscatter and finer sediments. Bathymetric contours on both plots, from the NOAA (2007) coastal relief model, are in meters.

RELATED PROJECTS

This work is not presently linked to any other programs, but could prove useful to ONR programs such as the Ripples DRI and the Shallow Water Acoustics '06 experiment, which will make use of interpolated point data related to seabed properties.

REFERENCES

Buczkowski, B.J., Reid, J.A., Jenkins, C.J., Reid, J.M., Williams, S.J., Flocks, J.G., 2006. usSEABED: Gulf of Mexico and Carribbean (Puerto Rico and U.S. Virgin Islands) Offshore Surficial Sediment Data Release. U.S. Geological Survey Data Series 146, version 1.0, Online at http://pubs.usgs.gov/ds/2006/146/.

Goff, J. A.,Wheatcroft, R. A., Lee, H., Drake, D. E., Swift, D. J. P., Fan, S., 2002. Spatial Variability of Shelf Sediments in the STRATAFORM Natural Laboratory, Northern California. Continental Shelf Research 22, 1199-1223.

Goff, J.A., Kraft, B.J., Mayer, L.A., Schock, S.G., Sommerfield, C.K., Olson, H.C., Gulick, S.P.S., Nordfjord, S., 2004. Seabed characterization on the New Jersey middle and outer shelf: Correlability and spatial variability of seafloor sediment properties. Marine Geology 209, 147-172.

Jenkins, C.J., 1997. Building Offshore Soils Databases. Sea Technology 38, 25-28.

Jenkins, C.J., 2002. Automated digital mapping of sediment colour descriptions. Geo-Marine Letters 22, 181-187.

NOAA, 2007. Coastal Relief Model. National Geophysical Data Center, Boulder, CO, USA. CD-ROM. Online at "http://www.ngdc.noaa.gov/mgg/coastal/coastal.html".

Reid, J.M., Reid, J.A., Jenkins, C.J., Hastings, M.E., Williams, S.J., Poppe, L.J, 2005. usSEABED: Atlantic coast offshore surficial sediment data release. U.S. Geological Survey Data Series 118, version 1.0. Online at http://pubs.usgs.gov/ds/2005/118/.

Reid, J.A., Reid, J.M., Jenkins, C.J., Zimmermann, M., Williams, S.J., Field, M.E., 2006. usSEABED: Pacific Coast (California, Oregon, Washington) Offshore Surficial-Sediment Data Release: U.S. Geological Survey Data Series 182, version 1.0. Available online at http://pubs.usgs.gov/ds/2006/182.

Schwab, W.C., Thieler, E.R., Allen, J.R., Foster, D.S., Swift, B.A., Denny, J.F., 2000. Influence of inner-continental shelf geologic framework on the evolution and behavior of the barrier-island system between Fire Island Inlet and Shinnecock Inlet, Long Island, New York. Journal of Coastal Research 16, 2, 408-422.

Williams, S. J., Jenkins, C., Currence, J., Penland, S., Reid, J., Flocks, J., Kindinger, J., Poppe, L., Kulp, M., Manheim, F., Hampton, M., Polloni, C., Rowland, J., 2003. New digital geological maps of U.S. continental margins: insights to sea floor sedimentary character, aggregate resources and processes. Proceedings of the International Conference on Coastal Sediments 2003, Corpus Christi, Texas, World Scientific Publishing Corporation and East Meets West Productions, Corpus Christi, Tex., 11 p., CD-ROM.

Williams, S.J., Reid, J.A., Arsenault, M.A., Jenkins, C., 2006. Characterization of sedimentary deposits using usSEABED for large-scale mapping, modeling and research of U.S. continental margins. Eos Transactions of AGU 87, Fall Meeting Supplement, Abstract H33B-1505.

PUBLICATIONS

Goff, J. A., C. Jenkins, B. Calder, 2006. Maximum likelihood resampling of noisy, spatially correlated data, Geochemistry, Geophysics and Geosystems, doi:10.1029/2006GC001297.
Goff, J. A., Jenkins, C.J., Williams, S.J. Characterization of seabed sediment variability using the usSEABED data base. Cont. Shelf Res., in review.
Jenkins, C. J., Goff, J. A. Competent interpolation for seabed substrates, with uncertainty calculations, Continental Shelf Research, in review.